MULTIMEDIA UNIVERSITY

STUDENT IDENTIFICATION NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 2, 2015/2016

BDS3024 – BUSINESS INTELLIGENCE

(All sections / Groups)

5 MARCH 2016
9.00 a.m – 11.00 a.m
(2 Hours)

INSTRUCTIONS TO STUDENT

1. This question paper consists of 5 pages with 5 questions only.
2. Attempt **ALL** questions. The distribution of the marks for each question is given.
3. Please write all your answers in the Answer Booklet provided.

QUESTION 1

A client is looking for a software company to develop customer relationship management system (CRMS). The client would like to use state-of-the-art technology to build the system. You as a system analyst suggest to the client to build the system on Apps. Describe the following terminologies to the client.

(a) data center. (5 marks)

(b) cloud computing. (5 marks)

(c) wi-fi. (5 marks)

(d) By using diagram(s), demonstrate how apps, wifi, cloud computing and data center are connected to each other to support the client in terms of CRMS. (11 marks)

(Total: 26 marks)

QUESTION 2

How can an expert system be used to detect probable fraud committed by a bank employee? (15 marks)

QUESTION 3

(a) What is auto-categorisation? (5 marks)

(b) How can auto-categorisation software help companies to serve customers and employees? (5 marks)

(c) Explain any **ONE** benefit that can be gained from this software. (5 marks)

(Total: 15 marks)

QUESTION 4

(a) What are data-mining techniques expected to find in the huge data warehouse? Provide **ONE** example. (8 marks)

(b) Data mining helps mainly in four ways: sequence analysis, classification, clustering, and forecasting. Data mining helps determine whether a person has committed fraud. Which of the four types of analysis help do that? Explain why. (7 marks)

(Total: 15 marks)

Continued...

CASE STUDY

A Safety Net

Many people would rather drive a car for many hours than fly. Perhaps this is because of the belief that if we have more control - such as when driving - we are also safer than if someone else - such as a pilot - maintains full control. Yet, statistics show clearly that air travel is the safest way to get from one point to another. This does not mean that government agencies do not try to improve air safety. They do. And they use data mining techniques to do so.

In 1950, 17 of every million commercial airline passengers were killed in airline accidents worldwide. As of the mid-1970s, with better airplanes but much more congested air traffic, the ratio has been close to one per million. In the United States, the ratio has been 0.3 air-traveller per million. Over the years, the Federal Aviation Administration (FAA) has collected and analyzed detailed records of accidents to learn from experience. Christopher Hart, the FAA systems administrator for System Safety, says that aviation is so safe now that if a single system errs, it is unlikely the event will threaten passenger safety because remedying measures can be taken.

Whenever crashes do occur, FAA investigators examine the event from every conceivable angle to determine what went wrong and devise ways to ensure it does not happen again. Ironically, the rare occurrences of accidents are the main reason why so little can be learned how. Predicting future disasters from ones that actually occurred is practically impossible. Therefore, the agency uses computer simulations to find potential hazards.

Many accidents are the result of a combination of events rather than a single one. The agency uses modeling software to analyze what might happen if certain weather conditions, load, and mechanical mishaps can cause an irreversible problem. However, to be able to perform such analyses the FAA must overcome two hurdles: combining data that is held by thousands of disparate bodies, and handling data that is unstructured.

Continued...

Potentially useful data is held by many national and private airlines, manufacturers, maintenance companies, air traffic control towers, trade associations, labor unions, and air forces. Data analysis would be easier if, for instance, the FAA could aggregate all the maintenance records of the Boeing 777 from all the airlines that operate it. But airlines do not like to share such information for fear of lawsuits or hurting their reputations. Even when relevant data is available, some of it is stored in forms that are not readily available for digital analysis.

In 1996 the FAA and its counterparts in some other countries established the Global Aviation Information Network (GAIN) to foster sharing of data. Aviation data is kept in two forms: digital (from flight data recorders) and textual. The latter consists of notes written by pilots and ground crews.

GAIN includes working groups that specialize in various areas. One of them is Analytic Methods and Tools Working Group, The group approached Megaputer Intelligence, Inc., a company in Bloomington, Indiana, that specializes in data mining software. The company promotes PolyAnalyst, the "most comprehensive data mining suite."

The tool uses the WordNet dictionary developed by the Cognitive Science Laboratory at Princeton University. For a given industry it assigns words to subject categories. It helps create, import, and manage taxonomies, and performs auto-categorization of text records according to those taxonomies. Taxonomies are categories and subcategories within given subjects. PolyAnalyst then creates a pictorial representation of the topics and the connections among them.

One module of PolyAnalyst, TextAnalyst, was specifically developed to scour text documents and "fish" for the useful information. It summarizes multiple documents, develops a tree-like topic structure, and performs free-form queries.

GAIN tested the tool for six weeks with reports from Southwest Airlines pilots that described abnormal occurrences during various flight phases, such as take off, ascending, cruising, descending, and landing. At Southwest, such data is stored in an

Continued...

Oracle database consisting of 63 structured columns and one column in which pilots can use up to 4,000 words of free text to describe the incident. That text was examined by a human analyst. The process was time-consuming and prone to error because it relied on the analyst's memory of details and possible links of incidents from multiple text pieces.

PolyAnalyst included in its analysis all the data, both from the structured fields and the text field. In the test, it did so with 2,000 database records and generated a graphic illustration of types of anomalies for each aircraft model used by the airline. To drill down and receive more details from particular pilot reports, users can click on a point on the graph that addresses an incident type.

The challenge now for GAIN and the FAA is to convince other airlines to allow the software to access their reports. The FAA is not trying to transfer airlines' data to its own database, maintain text analysis tools, or analyze safety data. Rather, GAIN is trying to work out a method for airlines to access the data of other airlines via a network, and to make the airlines aware of the value of tools such as PolyAnalyst. The trick is to enable the airlines to access all the data of all the airlines but leave the data untraceable to a particular airline. When this is accomplished and more airlines use such analysis tools, air travel will be even safer than it is today.

Sources: Robb, D., "Mining Data to Up Airline Safety," Datamation, February 4, 2005; www.megaputercom, 2005; Robb, D.,"Text Mining Tools Take on Unstructured Data," Computerworld(www.computerworldcom), June 21. 2004; Temin, T. Fl., "Taking Data to a Higher Plane," Government Computer News (www.gcn.corn), June 14, 2004.

QUESTION 5

(a) Comparing flight records to a speeding ticket, both contain structured and text fields. Why is the text field necessary? (10 marks)

Continued...

(b) The challenge in developing software for text analysis is to give the software the ability to understand:

(i) the meaning of the same word in two different contexts, and

(ii) different words (synonyms) that have the same meaning in a given context.

List **THREE** examples of a word that means two different things in the context of two different industries (such as police work and flight). (9 marks)

(c) Why is it so important that airlines share their flight record resources in trying to predict accidents? (10 marks)

(Total: 29 marks)

End of Page